

(19) 日本国特許庁 (J P)

(12) 特 許 公 報 (B 2)

(11) 特許番号

第2975808号

(45) 発行日 平成11年(1999)11月10日

(24) 登録日 平成11年(1999) 9月3日

(51) Int.Cl.<sup>8</sup>  
G 1 0 L 3/02 3 0 1  
3/00 5 6 1

F I  
G 1 0 L 3/02 3 0 1 B  
3/00 5 6 1 A

請求項の数1(全 3 頁)

(21) 出願番号 特願平5-129454  
(22) 出願日 平成5年(1993)5月31日  
(65) 公開番号 特開平6-337697  
(43) 公開日 平成6年(1994)12月6日  
審査請求日 平成9年(1997)4月24日

(73) 特許権者 000001889  
三洋電機株式会社  
大阪府守口市京阪本通2丁目5番5号  
(72) 発明者 井出 廣一  
大阪府守口市京阪本通2丁目18番地 三  
洋電機株式会社内  
(72) 発明者 渡辺 俊幸  
大阪府守口市京阪本通2丁目18番地 三  
洋電機株式会社内  
(74) 代理人 弁理士 安富 耕二 (外1名)

審査官 松尾 淳一

(56) 参考文献 特開 平1-236000 (J P, A)  
特開 昭61-180296 (J P, A)  
特公 昭62-57040 (J P, B 2)

最終頁に続く

(54) 【発明の名称】 音声認識装置

(57) 【特許請求の範囲】

【請求項1】 マイクロフォンと、このマイクロフォンから得られる音声信号を増幅する可変利得アンプと、このアンプから得られる出力を入力とするアナログ・デジタル変換器と、この変換器からの出力を処理する音声認識処理回路とを備え、前記音声認識処理回路における音声認識率に基づいて前記可変利得アンプの利得を調整するようにした音声認識装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は入力手段として音声を用いる情報機器や民生機器一般に利用できる音声認識装置に関する。

【0002】

【従来の技術】 人間の発声音は発声者ごとに異なると共に同一発声者でも発声のたびに変動する。初期の音声認識技術ではこれらの時間領域での変動をいかに整合させパターンマッチングするかの検討がなされた。さらに、数理統計手法やニューラルネットを導入する事で周波数・スペクトル両領域でのモデル化が行われ不特定話者の単語単位での認識率は格段に向上してきた。

【0003】 このような技術開発に伴い現在では音声認識が手軽に行えるパーソナルコンピュータ内蔵ボードが多数製品化されている。

【0004】 一方、民生品分野でも複雑化したAV機器の操作の単純化の1つのアプローチとして音声入力による操作入力という試みがなされており、例えば、雑誌「テレビ技術」( '91年5月号 第38頁~第44頁) には音声認識技術を利用したリモートコントロール

装置が紹介されているように一部商品化されている。

【0005】一般的な音声認識に使用される回路構成例を図2に示す。一般にはデジタル信号処理が行われるため、入力マイク1・アンプ系2とAD（アナログ・デジタル）変換処理部3とコンピュータやDSP（デジタル・シグナルプロセッサ）もしくは専用の音声認識用ICで構成される認識処理部4と外部インターフェイス部5に分けられる。

【0006】

【発明が解決しようとする課題】現在製品化されている技術は上述の単語単位（ユニット）あるいはその逐次処理の範疇のもので、この場合、予め決められた単語のみを認識できれば良いので切り出した音声と用意したデータベースをどれだけうまくマッチングさせるかで認識率が決まる。

【0007】これは、認識のアルゴリズム上の問題と共に周辺回路でどれほどうまく必要な音声を取り込めるのかも重要な要素となる。

【0008】簡単のため入力音声信号が図3に示すような単一周波数であるとして考える。入力が過大レベルの場合、認識の演算部へ取り込む前のアンプ系でオーバレンジとなって図4に示すとおり波形はピークがクリップされ台形状となり高調波成分が重畳する。この過大入力レベルでは波形自体が変わってしまっている。つまり、認識の重要なパラメータである周波数のパターンが変わってしまい認識できなくなってしまう。

【0009】上述の様な過大入力時の対応としてはレンジオーバとならない用に予めアンプゲインを落としてしまうことも考えられるが、逆に過小レベルの認識率が大幅に低下してしまう。また、入力レベルに応じてゲインを変えさせる方法もあるが、予めある程度入力レベルが決まっており同期信号等の基準がある映像信号と異なり音声信号はもともと時間軸・レベルとも基準がなく、かなりダイナミックに変動するため、この方法では危険である。

【0010】そこで、本考案では音声認識結果を使って入力部の調整を行うことで認識の精度を向上させようと言うものである。

【0011】

【課題を解決するための手段】同一人物がマイク等に向かって発生する場合、ある一定時間内に決められた単語を発生する時を考えると、マイクを比較的同じ様な位置に持ってゆき同じ様な音圧で発生することが多い。そこで、音声入力もレベルオーバ等で歪む場合は連続して発生すると考えられる。

【0012】そこで、音声認識に不成功の場合はゲイン調整をする。この時、学習効果の有るような手法例えば

過去数回のデータを平均するなどして記憶して置くようにする。

【0013】

【作用】このようにすることで上述のような発声の特性から認識率が向上していく。

【0014】

【実施例】図1に本発明の一実施例を示す。

【0015】マイク1から入力された音声はプリアンプ2A、可変利得アンプ2Bで増幅後、AD変換器3でデジタル信号に変換されDSP4に供給される。DSPでは音声認識処理を行った後インターフェイス回路5から外部機器に任意のフォーマットもしくはポートを通じて出力される。

【0016】ここで、認識が不成功の場合、DSP4からアンプゲインを調整する信号が出力され、この信号が可変利得アンプ2Bの利得切り換え回路2Cを切り換える。即ち、音声認識が不成功の場合は可変利得アンプ2Bの利得を増加せしめるように切り替える。

【0017】尚、認識の成功・不成功は確率的な問題なので、ゲイン調整については人間が結果をフィードバックしてやる方法やDSP側で確率的に学習させる方法等を使用するようにしても良い。

【0018】音声認識では先述の様に周波数成分が重要なパラメータとなっており、レンジオーバの様な特徴的な歪であれば周波数分析で有る程度の予測がつけられる場合もあり、アルゴリズムのプログラムROMに余裕があればこの情報も調整時の判別要素に加えればさらに精度は向上する。

【0019】

【発明の効果】本発明によれば音声認識の成功率の重要な要素となるマイクアンプ部での波形歪を認識結果をフィードバックすることで改善させたので、認識率を向上させることができる。

【図面の簡単な説明】

【図1】本発明を実施した音声認識装置のブロック図である。

【図2】従来の音声認識システムのブロック図である。

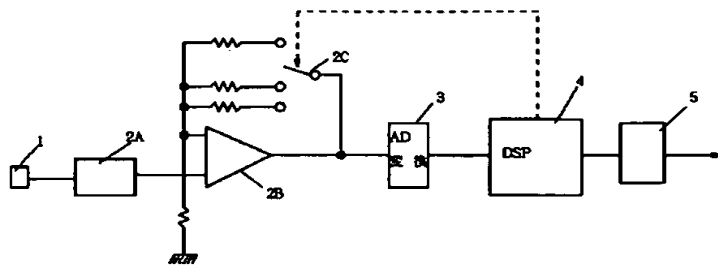
【図3】動作説明のための波形図である。

【図4】動作説明のための波形図である。

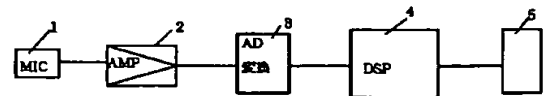
【符号の説明】

- 1    マイクロフォン
- 2A   プリアンプ
- 2B   可変利得増幅器
- 2C   利得切り換え回路
- 3    AD変換器
- 4    デジタル・シグナルプロセッサ
- 5    インターフェイス回路

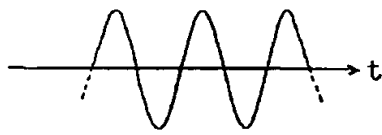
【図1】



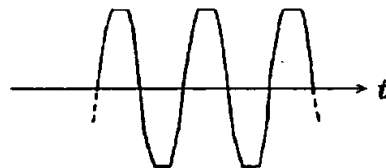
【図2】



【図3】



【図4】



フロントページの続き

(58)調査した分野(Int.Cl.<sup>6</sup>, DB名)

G10L 3/02 301

G10L 3/00 561

EPAT (QUESTEL)

INSPEC (DIALOG)

WPI (DIALOG)

Partial Translation of JP2975808

[0011]

[Means for Solving the Problems] Every time a person utters  
5 a decided word toward a microphone or the like in a certain  
constant time, he/she generally brings the microphone to a  
relatively similar position and utters the word with a similar  
sound pressure. Therefore, a voice input is conceivably also  
continuously generated when distorted due to an excessive level  
10 or the like.

[0012] Therefore, a gain is controlled if voice recognition  
is unsatisfactory. In this case, a technique having a learning  
effect is employed for averaging several precedent data and  
storing the same, for example.

15 [0013]

[Function] Thus, the recognition rate is improved due to the  
aforementioned characteristics of voice production.

[0014]

[Embodiment] Fig. 1 shows an embodiment of the present  
20 invention.

[0015] A voice input from a microphone 1 is amplified in a  
preamplifier 2A and a variable gain amplifier 2B, thereafter  
converted to a digital signal in an A-D converter 3 and supplied  
to a DSP 4. The digital signal is subjected to voice  
25 recognition processing in the DSP, and thereafter output from

an interface circuit 5 to an external device in an arbitrary format or through a port.

[0016] When recognition is unsatisfactory, The DSP 4 outputs a signal for controlling the amplifier gain, for switching a gain switching circuit 2C of the variable gain amplifier 2B. In other words, the gain switching circuit 2C is switched to increase the gain of the variable gain amplifier 2B when voice recognition is unsatisfactory.



## PATENT ABSTRACTS OF JAPAN

(11) Publication number: **06337697 A**(43) Date of publication of application: **06.12.94**

(51) Int. Cl.

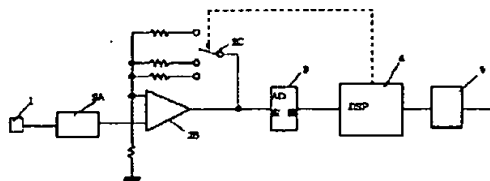
**G10L 3/02**(21) Application number: **05129454**(22) Date of filing: **31.05.93**(71) Applicant: **SANYO ELECTRIC CO LTD**(72) Inventor: **IDE KOICHI  
WATANABE TOSHIYUKI**(54) **VOICE RECOGNITION DEVICE**

COPYRIGHT: (C)1994,JPO

(57) Abstract:

**PURPOSE:** To improve the precision in recognition by using a voice recognition result and adjusting an input part.

**CONSTITUTION:** After a sound inputted from a microphone 1 is amplified by a preamplifier 2A and a variable gain amplifier 2B, is converted to a digital signal by an A/D converter 3 to be supplied to a digital signal processor(DSP) 4. After the digital signal is voice recognition processed in the DSP 4, is outputted from an interface circuit 5 through an optimum format or port. At this time, when no recognition succeeds, a signal adjusting an amplifier gain is outputted from the DSP 4, and a gain switching circuit 2C of the variable gain amplifier 2B is switched by the signal. That is, when no recognition succeeds, the circuit 2C is switched so as to increase the gain of the variable gain amplifier 2B. Further, since a success/a failure in recognition is a probable problem, a method feeding back the result by a human and the method, etc., probably learning in the DSP side 4 may be used also.



(19) 日本国特許庁 (J P)

(12) 特 許 公 報 (B 2)

(11) 特許番号

第2975808号

(45) 発行日 平成11年(1999)11月10日

(24) 登録日 平成11年(1999)9月3日

(51) Int.Cl. <sup>8</sup>	識別記号	F I
G 1 0 L 3/02	3 0 1	G 1 0 L 3/02 3 0 1 B
3/00	5 6 1	3/00 5 6 1 A

請求項の数1(全 3 頁)

(21) 出願番号	特願平5-129454	(73) 特許権者	000001889 三洋電機株式会社 大阪府守口市京阪本通2丁目5番5号
(22) 出願日	平成5年(1993)5月31日	(72) 発明者	井出 廣一 大阪府守口市京阪本通2丁目18番地 三 洋電機株式会社内
(65) 公開番号	特開平6-337697	(72) 発明者	渡辺 俊幸 大阪府守口市京阪本通2丁目18番地 三 洋電機株式会社内
(43) 公開日	平成6年(1994)12月6日	(74) 代理人	弁理士 安富 耕二 (外1名)
審査請求日	平成9年(1997)4月24日	審査官	松尾 淳一
		(56) 参考文献	特開 平1-236000 (J P, A) 特開 昭61-180296 (J P, A) 特公 昭62-57040 (J P, B 2)

最終頁に続く

(54) 【発明の名称】 音声認識装置

1

(57) 【特許請求の範囲】

【請求項1】 マイクロフォンと、このマイクロフォンから得られる音声信号を増幅する可変利得アンプと、このアンプから得られる出力を入力とするアナログ・デジタル変換器と、この変換器からの出力を処理する音声認識処理回路とを備え、前記音声認識処理回路における音声認識率に基づいて前記可変利得アンプの利得を調整するようにした音声認識装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は入力手段として音声を用いる情報機器や民生機器一般に利用できる音声認識装置に関する。

【0002】

2

【従来の技術】 人間の発声音は発声者ごとに異なると共に同一発声者でも発声のたびに変動する。初期の音声認識技術ではこれらの時間領域での変動をいかに整合させパターンマッチングするかの検討がなされた。さらに、数理統計手法やニューラルネットを導入する事で周波数・スペクトル両領域でのモデル化が行われ不特定話者の単語単位での認識率は格段に向上してきた。

【0003】 このような技術開発に伴い現在では音声認識が手軽に行えるパーソナルコンピュータ内蔵ボードが多数製品化されている。

【0004】 一方、民生品分野でも複雑化したAV機器の操作の単純化の1つのアプローチとして音声入力による操作入力という試みがなされており、例えば、雑誌「テレビ技術」( '91年5月号 第38頁〜第44頁) には音声認識技術を利用したリモートコントロール

装置が紹介されているように一部商品化されている。

【0005】一般的な音声認識に使用される回路構成例を図2に示す。一般にはデジタル信号処理が行われるため、入力マイク1・アンプ系2とAD（アナログ・デジタル）変換処理部3とコンピュータやDSP（デジタル・シグナルプロセッサ）もしくは専用の音声認識用ICで構成される認識処理部4と外部インターフェイス部5に分けられる。

【0006】

【発明が解決しようとする課題】現在製品化されている技術は上述の単語単位（ユニット）あるいはその逐次処理の範疇のもので、この場合、予め決められた単語のみを認識できれば良いので切り出した音声と用意したデータベースをどれだけうまくマッチングさせるかで認識率が決まる。

【0007】これは、認識のアルゴリズム上の問題と共に周辺回路でどれほどうまく必要な音声を取り込めるのかも重要な要素となる。

【0008】簡単のため入力音声信号が図3に示すような単一周波数であるとして考える。入力が過大レベルの場合、認識の演算部へ取り込む前のアンプ系でオーバレンジとなって図4に示すとおり波形はピークがクリップされ台形状となり高調波成分が重畳する。この過大入力レベルでは波形自体が変わってしまっている。つまり、認識の重要なパラメータである周波数のパターンが変わってしまい認識できなくなってしまう。

【0009】上述の様な過大入力時の対応としてはレンジオーバーとならない用に予めアンプゲインを落としてしまうことも考えられるが、逆に過小レベルの認識率が大幅に低下してしまう。また、入力レベルに応じてゲインを可変させる方法もあるが、予めある程度入力レベルが決まっており同期信号等の基準がある映像信号と異なり音声信号はもともと時間軸・レベルとも基準がなく、かなりダイナミックに変動するため、この方法では危険である。

【0010】そこで、本考案では音声認識結果を使って入力部の調整を行うことで認識の精度を向上させようと言うものである。

【0011】

【課題を解決するための手段】同一人物がマイク等に向かって発生する場合、ある一定時間内に決められた単語を発生する時を考えると、マイクを比較的同じ様な位置に持ってゆき同じ様な音圧で発生することが多い。そこで、音声入力もレベルオーバー等で歪む場合は連続して発生すると考えられる。

【0012】そこで、音声認識に不成功の場合はゲイン調整をする。この時、学習効果の有るような手法例えば

過去数回のデータを平均するなどして記憶して置くようにする。

【0013】

【作用】このようにすることで上述のような発声の特性から認識率が向上していく。

【0014】

【実施例】図1に本発明の一実施例を示す。

【0015】マイク1から入力された音声はブリアンプ2A、可変利得アンプ2Bで増幅後、AD変換器3でデジタル信号に変換されDSP4に供給される。DSPでは音声認識処理を行った後インターフェイス回路5から外部機器に任意のフォーマットもしくはポートを通じて出力される。

【0016】ここで、認識が不成功の場合、DSP4からアンプゲインを調整する信号が出力され、この信号が可変利得アンプ2Bの利得切り換え回路2Cを切り換える。即ち、音声認識が不成功の場合は可変利得アンプ2Bの利得を増加せしめるように切り替える。

【0017】尚、認識の成功・不成功は確率的な問題なので、ゲイン調整については人間が結果をフィードバックしてやる方法やDSP側で確率的に学習させる方法等を使用するようにしても良い。

【0018】音声認識では先述の様に周波数成分が重要なパラメータとなっており、レンジオーバーの様な特徴的な歪であれば周波数分析で有る程度の予測がつけられる場合もあり、アルゴリズムのプログラムROMに余裕があればこの情報も調整時の判別要素に加えればさらに精度は向上する。

【0019】

【発明の効果】本発明によれば音声認識の成功率の重要な要素となるマイクアンプ部での波形歪を認識結果をフィードバックすることで改善させたので、認識率を向上させることができる。

【図面の簡単な説明】

【図1】本発明を実施した音声認識装置のブロック図である。

【図2】従来の音声認識システムのブロック図である。

【図3】動作説明のための波形図である。

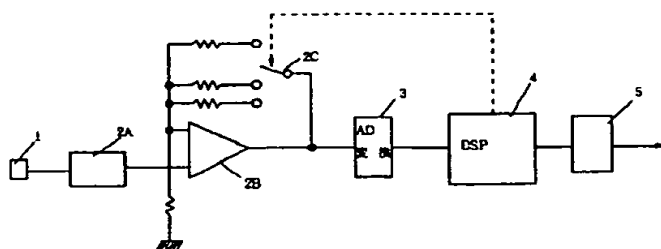
【図4】動作説明のための波形図である。

【符号の説明】

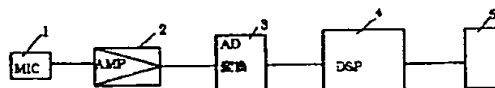
- 1    マイクロフォン
- 2A   ブリアンプ
- 2B   可変利得増幅器
- 2C   利得切り換え回路
- 3    AD変換器
- 4    デジタル・シグナルプロセッサ
- 5    インターフェイス回路



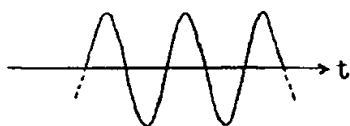
【図1】



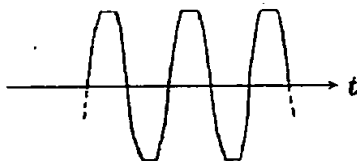
【図2】



【図3】



【図4】



フロントページの続き

(58)調査した分野(Int.Cl.<sup>6</sup>, DB名)

G10L 3/02 301

G10L 3/00 561

EPAT (QUESTEL)

INSPEC (DIALOG)

WPI (DIALOG)